

Centre for Economic Growth and Policy

Two Little Treasure Games driven by Unconditional Regret

Vladislav Damjanovic

Working Paper 2016 (06)

Two “Little Treasure Games” driven by Unconditional Regret

Vladislav Damjanovic*

Durham University Business School

October 18, 2016

Abstract

For Traveler’s Dilemma and Minimal Effort Coordination games, the unconditional regret matching (URM) procedure predicts outcomes close to the experimental ones. This supports a claim that the URM procedure can be well suited to predict the behavior of experimental subjects in repeated games.

JEL Classification: D8 C7 D7

Keywords: Adaptive Heuristics, Regret Matching, Coarse-Correlated Equilibrium, Traveler’s Dilemma Game, Minimal Effort Coordination Game

1 Introduction

We argue that the behavior of experimental subjects in a repeated game can be viewed as a consequence of following the URM procedure proposed in Hart and Mas-Colell (2000, 2001), hereafter HM¹.

We implement URM in a repeated Traveler’s Dilemma game (TDG) originating in Basu (1994) and a Minimum Effort Coordination game (MECG) introduced in Van Huyck et al. (1990) to match the experimental outcomes in Capra et al. (1999) (CGGH) and Goeree and Holt (2005) (GH). Both games belong to the “little treasures of game theory” (Goeree and Holt, 2001) and they have been extensively studied in the literature; see Anderson et al. (2001), Rubinstein (2006) and Eichberger and Kelsey (2011) among others. In Goeree and Holt (2001), experiments show that outcomes can be highly sensitive to the change in payoff structure which can produce outcomes at odds with the prediction of the Nash equilibrium. Standard explanations of deviations from the Nash equilibrium are the Quantal

*vladislav.damjanovic@durham.ac.uk

¹See also Young (2007).

Response Equilibrium (McKelvey and Palfrey, 1995) and k-level bounded rationality (Stahl and Wilson, 1995) among others. Eichberger and Kelsey (2011) show that the experimental outcomes of treasure games may be a consequence of experimental subjects perceiving opponents' behavior as ambiguous.

We show that in the short run, the URM procedure predicts outcomes close to the experimental ones evidenced in CGGH and GH. In case of TDG, the game departure from a Nash equilibrium is temporary since in the long run, and irrespective of cost, the joint play converges to the unique Nash equilibrium. In the case of the MECG, the joint play converges to Nash equilibria that are inversely related to the cost parameter. We also discuss dynamics of joint play towards equilibrium induced by URM with bounded memory (see Saran and Serrano, 2013).

2 Regret Matching

Following HM (2001), consider a particular player i from a group of players $N = \{1, 2, \dots, n\}$ engaged in a game against her opponent(s), playing an action $s_t^i \in S$ at time t . Let $\bar{u}_t^i = \frac{1}{t} \sum_{t' \leq t} u_{t'}^i(s_{t'}^i, s_{t'}^{-i})$ be the realized average payoff she received up to time t given the history $h_t^i := ((s_{t'}^i, s_{t'}^{-i}), t' = 1, 2, \dots, t)$ which is a collection of her own actions and the actions of her opponents denoted by $s_{t'}^{-i}$. Let $\bar{u}_t^i(s) = \frac{1}{t} \sum_{t' \leq t} u_{t'}^i(s, s_{t'}^{-i})$ be the average payoff she would receive had she played constant action s in all periods $t' \leq t$ and all other players played as they did. Then

$$r_t(s) = \bar{u}_t(s) - \bar{u}_t^i \quad (1)$$

is her unconditional regret of not playing action s . URM prescribes playing each action s in the next period with a probability proportional to the positive part of its unconditional regret:

$$p_{t+1}(s) = [r_t(s)]_+ / \sum_{s'} [r_t(s')]_+ . \quad (2)$$

In the long run, URM leads to no-regret for all players, and their joint actions converge to the set of coarse-correlated equilibria CCE (Moulin and Vial, 1978; Hart and Mas-Colell, 2000; Young, 2004). For a two-player game, CCE is a joint probability distribution $P = (p_{s,s'})$ over $S^1 \times S^2$ if

$$\begin{aligned} \sum_{(s^1, s^2) \in S^1 \times S^2} u_1(s^1, s^2) p_{s^1 s^2} &\geq \sum_{(s^1, s^2) \in S^1 \times S^2} u_1(x^1, s^2) p_{s^1 s^2}, \text{ for all } x^1 \in S^1 \\ \sum_{(s^1, s^2) \in S^1 \times S^2} u_2(s^1, s^2) p_{s^1 s^2} &\geq \sum_{(s^1, s^2) \in S^1 \times S^2} u_2(s^1, y^2) p_{s^1 s^2}, \text{ for all } y^2 \in S^2. \end{aligned} \quad (3)$$

3 TD Game

Similar to CGGH, the actions are integers from 80 to 200, i.e. $s^i \in S := \{80, 81, \dots, 199, 200\}$ and $N = \{1, 2, \dots, 10\}$. The i th player's payoff function is:

$$u^i(s^i, s^j) = \min(s^i, s^j) + C \cdot \text{sign}(s^j - s^i), \quad (4)$$

with $i, j \in N, i \neq j$. For any value of punishment/reward parameter $C > 1$ claiming the minimal amount, $s^* = 80$ is a unique Nash equilibrium. The experimental outcomes (CGGH) presented in Figure 1 show an inverse relationship between average claim and parameter C .

The sequences of players' claims in period t , $\{s_t^1, s_t^2, \dots, s_t^{10}\}$ are generated by URM and the average claim is $s_t = (1/n) \sum_{i=1}^n s_t^i$, to which we refer as the claim in period t . Initially, $s_0^i, i = 1, 2, \dots, 10$ are uniformly drawn from S . In each subsequent round, $t = 1, 2, \dots, T$ players are randomly (uniformly) matched and their payoffs are calculated by (4). Then, given a history h_t^i , action $s_{t+1}^i \in S$ is played with a probability prescribed by (2). The expected claims averaged over $N = 1000$ simulations for 10 rounds exhibit an inverse relationship with C as shown in Figure 1.

For intermediary values of the punishment parameter C , the last period claims reported in CGGH are inside 95% confidence intervals. As C increases, the confidence intervals move downward, indicating an inverse relationship between the average claim and C . To produce a better match with experimental outcomes, one can use generalized regrets (Hart and Mas-Colell, 2001) to transform positive regrets differently across players and actions.

The convergence to equilibrium is illustrated by averaging over N simulated claims² at period T for each $C = 5 : 5 : 80$. As shown by simulations (left-hand panel in Figure 2), the expected claims are not only decreasing in costs but they also decline with T and converge to the unique Nash equilibrium $s^* = 80$. Moreover, for each value of C , the claim per period rests at the Nash equilibrium in a finite time T inversely related to C .

Evolutions of expected claims (averaged over 200 simulated paths) for different C (Figure 2) show that: (i) for all t , there is a strict ordering of average paths preserving the inverse relationship between the expected claims and punishment C and (ii) there is a critical C^* so that all paths corresponding to $C < C^*$ are non-monotonic, and all paths with $C \geq C^*$ are monotonically decreasing.

4 Minimum-Effort Coordination Game

In MECG, $N = \{1, 2, \dots, 10\}$ and actions (efforts) of player i , $s^i \in S := \{110, 111, \dots, 169, 170\}$ are chosen as in GH. The payoff function is:

²Initially, N is chosen to be larger to obtain a smoother distribution of claims.

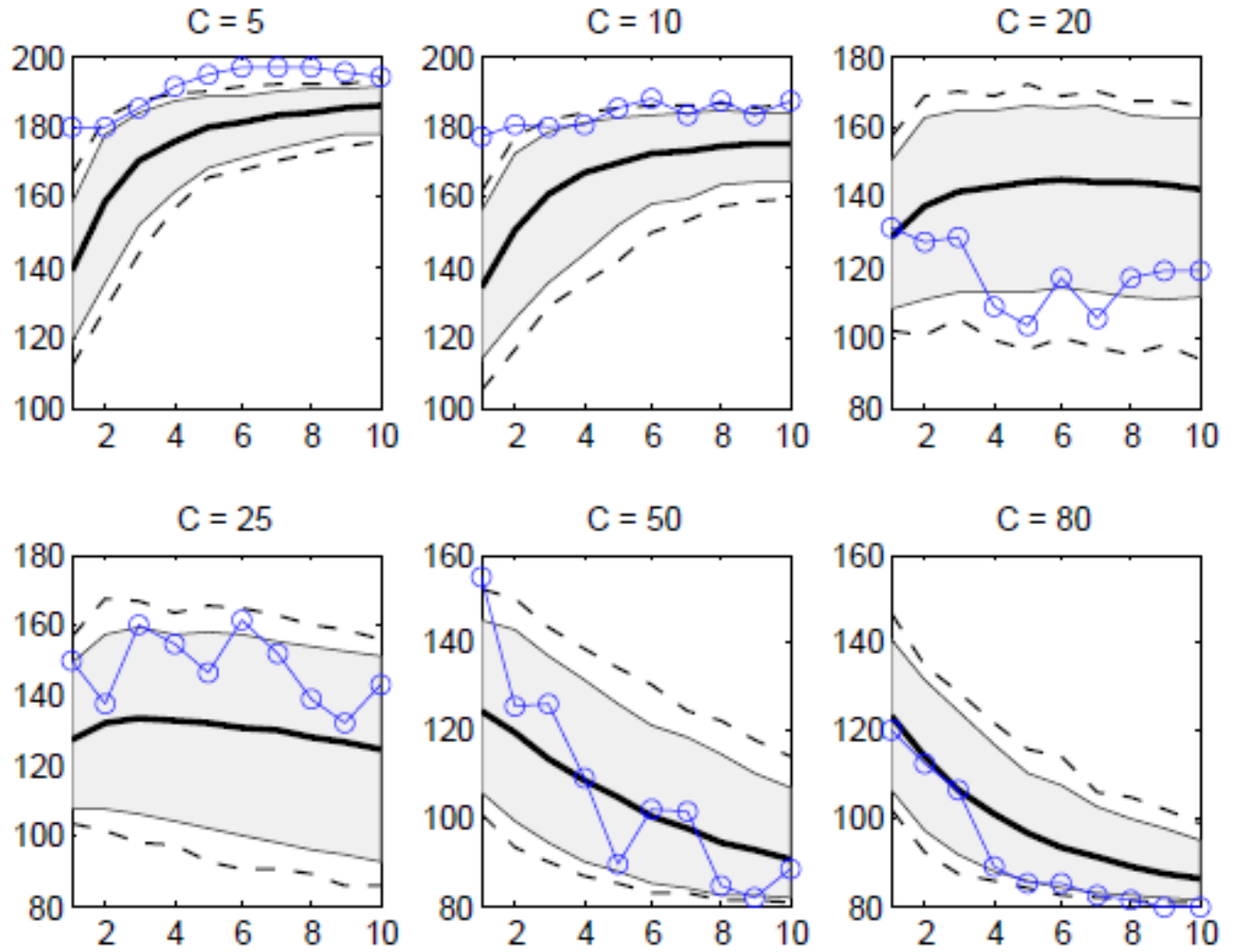


Figure 1: Expected claims (bold line) for different costs C with 95% (grey area) and 99% (between dashed lines) confidence intervals. Data from CCGH are presented by circles.

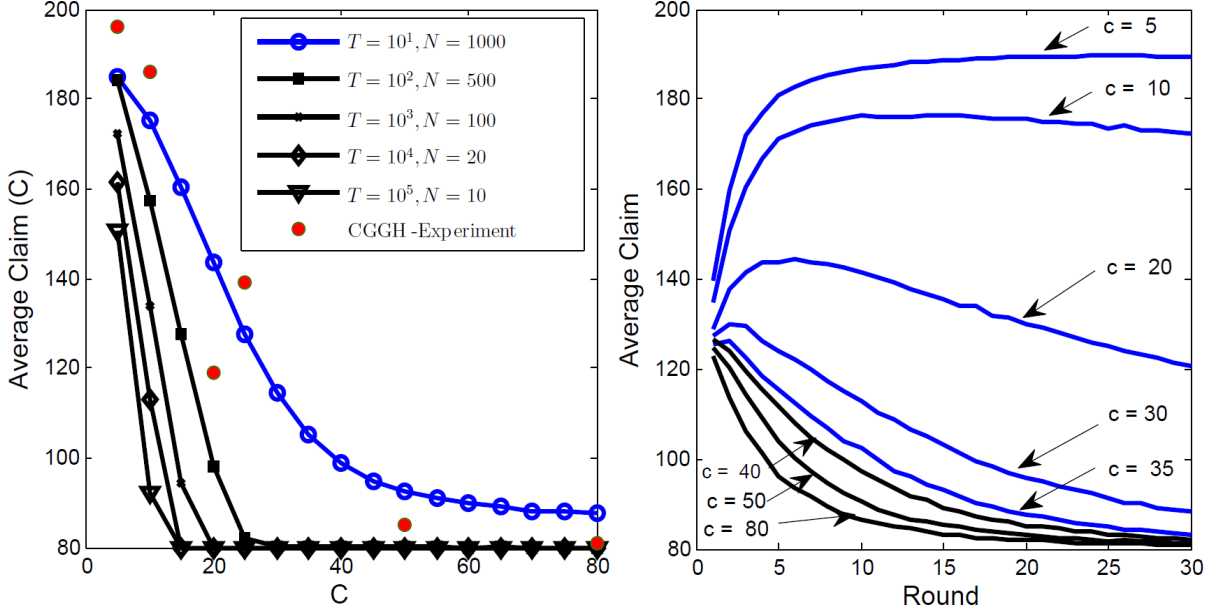


Figure 2: Average claims in period T and their evolutions in time.

$$u^i(s^i, s^j) = \min(s^i, s^j) - cs^i, \quad (5)$$

with $i, j \in N$, and the cost per unit of effort $c < 1$. In this game, any common effort s is a Nash equilibrium.

We focus on the dynamics of effort in round t , defined as $s_t = (1/N) \sum_{i=1}^N s_t^i$. The initial efforts are uniformly drawn from S . In round $t + 1$, given histories h_t^i , actions $s_{t+1}^i \in S$ are played with probabilities proportional to their unconditional regrets calculated by (1) and (5).

In Figure 3, the expected efforts for low and high values of c are presented together with experimental ones (GH) averaged across three different laboratory sessions. In line with the risk-dominance criterion,³ there is an inverse relationship between equilibrium efforts and costs. In equilibrium, all players coordinate on the same action after some rounds as shown in the left-hand panel of Figure 4.

To learn more about convergence to equilibrium and equilibrium selection, we calculate an average effort for $N = 100$ simulations for different c at periods $T = 10, 500, 1000$ (right-hand panel of Figure 4). Convergence to the limiting average effort is clearly established as the difference between average efforts in $T = 500$ (red line) and $T = 1000$ is negligible. The equilibrium effort is a strictly decreasing function of the cost parameter in some range of intermediary values of c . The highest effort is selected for some range of the lowest

³A unilateral increase for $e = 1$ will lead to a decrease in the payoff for the amount c while a unilateral decrease for $e = 1$ will lead to a decrease in the payoff for the amount $1 - c$. Like in a 2x2 game when $c < (>) 1/2$, choosing a higher (lower) effort is a risk dominant action (see HG, 2005).

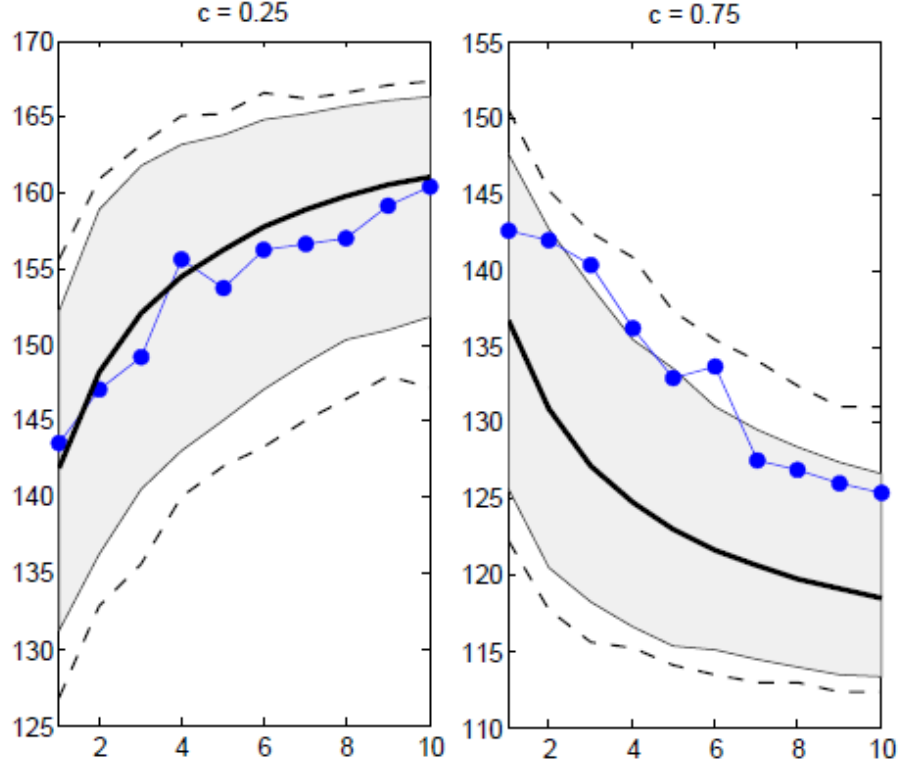


Figure 3: Expected efforts (averaged over 1000 simulations) during the first ten periods with 95% (grey area) and 99% (dashed lines) confidence intervals. Experimental data from GH are presented by dots.

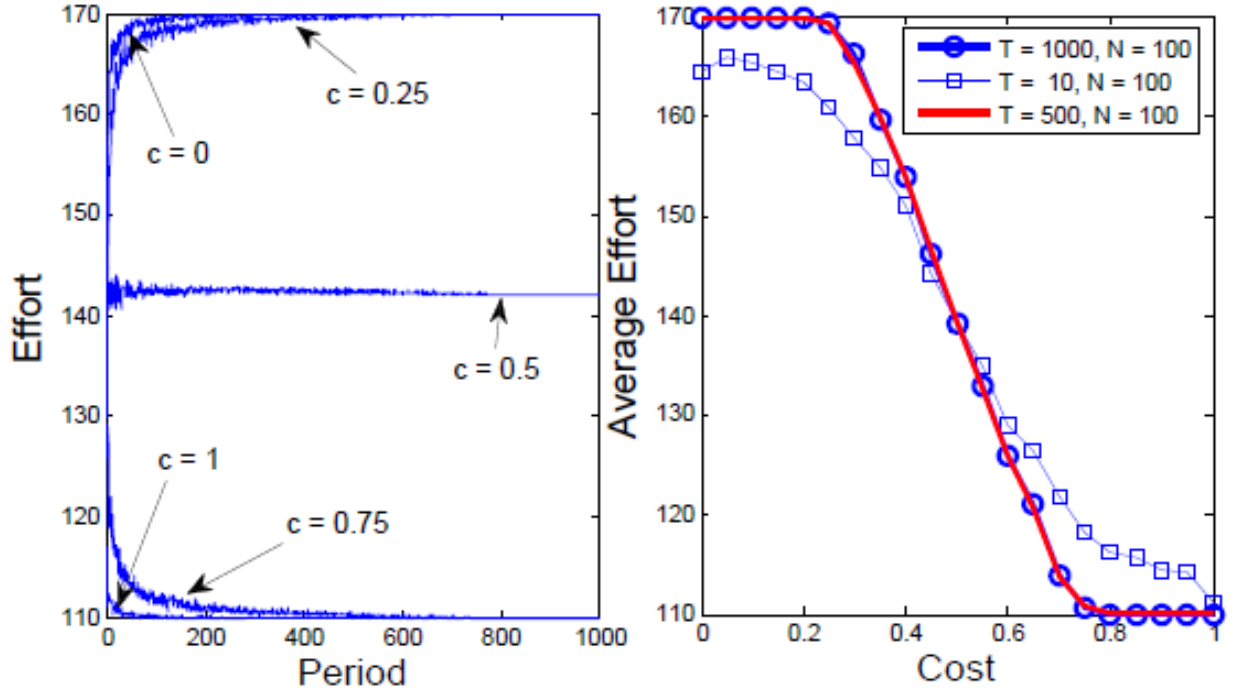


Figure 4: Evolutions of efforts and convergence to equilibria.

values of c , while the lowest effort is selected for some range of highest cost values, which partly conforms with the equilibrium prediction based on maximization of potential⁴ which is supported in an experiment by Van Hyck et al. (1990).

5 URM with Finite Memory

As in Saran and Serrano (2012), we introduce bounded memory in the player's strategy assuming that she remembers the last $m \geq 1$ rounds, so that (1) becomes

$$r_t^m(s) = \frac{1}{m} \sum_{t'=t-m+1}^t (u_{t'}^i(s, s_{t'}^{-i}) - u_{t'}^i(s_{t'}^i, s_{t'}^{-i})) \text{ for all } t \geq m \quad (6)$$

with the initial history being generated by URM

$$r_t^m(s) = \frac{1}{t} \sum_{t'=1}^t (u_{t'}^i(s, s_{t'}^{-i}) - u_{t'}^i(s_{t'}^i, s_{t'}^{-i})) \text{ for all } t < m. \quad (7)$$

Evolutions of expected actions (averaged over $N = 200$) for both games with different initial conditions, costs, and memories are shown in Figure 5. To check for the robustness of our results with respect to the initial conditions, we choose the initial distribution of claims $\{3, [80, \dots, 194]; 7, [195, \dots, 200]\}$ by uniformly drawing 3 claims from $[80, \dots, 194]$ and 7 claims from $[195, \dots, 200]$ ⁵. The bold black line in the left-hand panel provides a simulation for TDG with $C = 5$.

Expected claims with finite memories $m = 100, 50, 20$ converge faster to the Nash as compared to the unbounded memory case. The shorter the memory, the faster is the convergence. However, by further decreasing the memory, the memory would, at some point, be insufficient to learn Nash and the dynamics could exhibit cycle-like behavior as for $m = 9$ (red line). The exemption is $m = 1$, where we have convergence to the Nash⁶.

Simulations show that for small C initially the learning process goes towards higher claims until the information content (history) is sufficient for URM to learn 'right' direction. Necessary conditions for that to happen are that the initial claims should not be 'too close' in action space and that the memory should not be too short, i.e. $m > 1$. For example, for degenerate distributions $s_0^i = 180, i = 1, \dots, 10$, and distributions with initial claims in a small vicinity of either boundary, the non-monotonicity disappears. Further simulations indicate that the non-monotonicity of average claims for small costs is robust with respect to the wide range of initial conditions.

⁴Maximization of the potential will select 170 for all $c < 1/2$, 110 for all $c > 1/2$, while for $c = 1/2$, any equilibrium effort can be selected (Monderer and Shapley, 1996).

⁵In Rubinstein (2005), aggregated data show that in one-shot TDG, 70% of the claims were in close vicinity of the maximal claim.

⁶It can be shown that when $m = 1$, the average claims decrease: $\bar{s}_t \leq \bar{s}_{t-1}$.

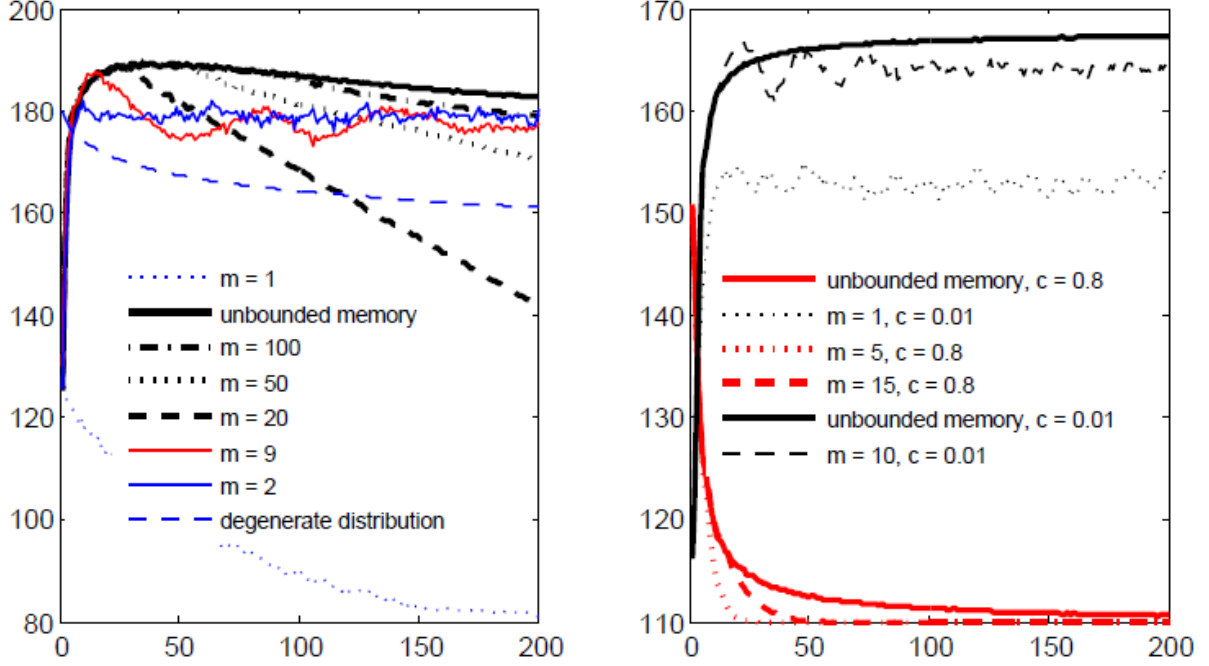


Figure 5: Evolutions of expected claims ($C = 5$) and expected efforts ($c = 0.01, 0.08$)

For MECG, the expected efforts are shown in the right-hand panel of Figure 5. Here, shortening the memory does not accelerate the convergence to the Nash as shown for the distribution of initial efforts $\{10, [110, 111]\}$ with $c = 0.01$ and memories $m = 1, 10$ (black lines) as compared to the unbounded memory case (black bold line). However, for $c = 0.8$ and initial distribution of claims $\{2, [110, \dots, 115]; 8, [165, 170]\}$, shortening the memory $m = 5, 15$ (red broken lines) accelerates the convergence to the Nash.

Our simulations suggest that URM with a finite memory may converge to a CCE different from Nash. For related results with finite memory in conditional regret matching, see Saran and Serrano (2006). It would be desirable to get a full characterization of CCEs. Below, we describe one subset of CCE for TDG.

Proposition 1 *In a TD game with $B - A \geq C > 1$, the following joint probability distribution represents the CCE equilibrium*

$$p_{zz} = p; \quad p_{yy} = 1 - p.$$

where $y \geq z + C$; $(1 - p)(y - z) \geq C - 1$; $p \geq \frac{C-1}{2C-1}$;

Proof. Following (3), we need to show that when Player 1 commits to follow the mediator's advice before it has been seen, the unconditional expectation of her payoff is at least as high as her expected payoffs from playing any other strategy provided that her opponent commits.

Since the game is symmetric, we consider one player with the unconditional expected payoff $U^{CC} = pz + (1 - p)y$. i) Consider any strategy x , such that $x < z$. The expected pay off is $x + C$, and to ensure that this strategy is not preferable, it must be $pz + (1 - p)y > z - 1 + C \geq x + C \Leftrightarrow (1 - p)(y - z) > C - 1$. ii) For $x = z$; the payoff is $pz + (1 - p)(z + C)$, and the necessary condition is $pz + (1 - p)y > pz + (1 - p)(z + C) \Leftrightarrow y \geq z + C$. iii) if $z < x < y$, the necessary condition is $pz + (1 - p)y > p(z - C) + (1 - p)(y - 1 + C) \Leftrightarrow p \geq \frac{C-1}{2C-1}$; iv) if $x \geq y$, the necessary condition is always satisfied. ■

The CCE may deliver a higher expected payoff for a lower C . The maximum $U^{CC} = B - C + 1 + \frac{C-1}{2C-1}$ is achieved when $y = B$, $z = B - 2C + 2$, $p = \frac{C-1}{2C-1}$ and it declines in C .

6 Conclusion

In our setting, agents learn to play the Nash equilibrium and in contrast to CGGH and GH, deviations from rationality are temporary. This is expected since TDG is a dominance solvable and MECG is a potential game and, as pointed out in Hart (2005), there exist adaptive heuristics that lead to Nash equilibrium. The initial conditions do not affect the equilibrium outcome, but can affect the dynamics towards equilibrium.

Our simulations do also show that the nature of the equilibrium outcome in games with finite memory URM could be different from the case with infinite histories. Shortening memory may lead to Pareto improvement as in TDG. This should be further investigated.

7 Acknowledgment

I am grateful to Andrew Caplin, Gareth Myles, Ariel Rubinstein, Larry Samuelson, and Andrew Schotter, as well as seminar participants at Durham University and University of St Andrews for their comments. Financial support from the Carnegie Trust for the Universities of Scotland (Project: Adaptive Heuristics in Repeated Games, 2009) is acknowledged. The usual disclaimer applies.

References

- [1] Anderson, S.P. , Goeree, J.K., Holt, C.A., 2001. Minimum-Effort Coordination Games: Stochastic Potential and Logit Equilibrium. *Games Econ. Behav.* 34(2), 177-199.
- [2] Basu, K., 1984. The Traveler's Dilemma: Paradoxes of Rationality in Game Theory. *American Econ. Rev.* 84(2), 391-395.
- [3] Capra, C. M., Goeree, J. K., Gomez, R., Holt, C. A. 1999. Anomalous Behavior in a Traveler's Dilemma?. *American Econ. Rev.* 89(3), 678-690.

- [4] Eichberger J., David Kelsey D., 2011, Are the treasures of game theory ambiguous? *Econ. Theory.* 48, 313–339.
- [5] Goeree, K.J., Holt A.C., 2005. An experimental study of costly coordination. *Games Econ. Behav.* 51(2), 349-364.
- [6] Goeree, K.J., Holt A.C., 2001. Ten Little Treasures of Game Theory and Ten Intuitive Contradictions. *American Econ. Rev.* 91(5), 1402 - 1422.
- [7] Hart S., Mas-Colell A., 2001. A general class of Adaptive strategies. *Journal of Econ. Theory.* 98, 26-54.
- [8] Hart S., Mas-Colell A., 2000. A Simple Adaptive Procedure Leading to Correlated Equilibrium. *Econometrica.* 68(5), 1127-1150.
- [9] McKelvey R.D., Palfrey, T.R., 1995. Quantal Response Equilibria for normal forms game. *Games Econ. Behav.* 10, 6 - 38.
- [10] Moulin H., Vial J. P., 1978. Strategically zero-sum games: The class of games whose completely mixed equilibria cannot be improved upon. *Intern. Journal.Game Theory.* 7 (3/4), 201-221.
- [11] Monderer, D., Shapley L. S. 1996. Potential Games, *Games Econ. Beh.* 14, 124-141.
- [12] Rubinstein A., 2006. Dilemmas of an Economic Theorist. *Econometrica.* 74(4), 865-883.
- [13] Saran R., Serrano R., 2012. Regret matching with finite memory. *Dyn Games Appl.* 2: 160-175.
- [14] Stahl O.D., Wilson W.P., 1995. On the Players' Models of Other Players: Theory and Experimental Evidence., *Games Econ. Behav.* 10, 218 - 254.
- [15] Van Huyck B.J., Battalio C.R., Richard O. B., 1990. '. *American Econ. Rev.* 80, 234-248.
- [16] Young H.P., 2007. The Possible and the Impossible in Multi-Agent Learning. *Artificial Intelligence.* 171, 429-433.
- [17] Young H.P., 2004. *Strategic Learning and its limits.* Oxford University Press